
THE APPLE OF KANT'S ETHICS: I-MAXIMS AS THE LOCUS OF ASSESSMENT

BY

SAMUEL KAHN

Abstract: I want to distinguish between maxims at three levels of abstraction. At the first level are what I shall call individual maxims, or i-maxims: maxim tokens as adopted by particular rational beings. At the second level are abstract maxims, or a-maxims: abstract principles distinct from any individual who adopts them. At the third level are maxim kinds, or k-maxims: sets of various action-guiding principles that are grouped on the basis of their content. In this paper, I argue for the thesis that i-maxims are the locus of assessment in Kant's ethics.

In this paper, I argue that what I call i-maxims are the locus of assessment in Kant's ethics. The paper is divided into four sections. In Section 1, I explain my thesis. In Section 2, I introduce textual support for my thesis. In Section 3, I introduce philosophical support for my thesis. In Section 4, I consider some objections.

1. Explaining my thesis

A maxim is a subjective principle of volition. But I want to distinguish between maxims at three levels of abstraction. At the first level are what I shall call individual maxims, or i-maxims. These are maxims as they are adopted by actual, concrete individuals. For instance, consider the maxim, 'to make a lying promise in order to get some ready money.' If two individuals, X and

Y, adopt this maxim, then, by virtue of being distinct individuals, they have numerically distinct even though qualitatively identical i-maxims.

But X and Y clearly have adopted a single maxim between them in some sense. This is the second level of abstraction. It requires one to take a maxim to be an abstract principle distinct from any individual who adopts it. I shall call these abstract maxims, or a-maxims. The two individuals in the previous paragraph, X and Y, have the same a-maxim but distinct i-maxims.

At the third level of abstraction are what might be called maxim kinds. This is when various maxims are grouped together, usually on the basis of their content. For instance, suppose that Z adopts the i-maxim, 'to make a lying promise in order to get a car.' This i-maxim is qualitatively similar to although nonetheless distinct from the i-maxims adopted by X and Y. These i-maxims might be grouped together into maxim kinds, or k-maxims. The question that arises now is: should the locus of assessment in Kant's ethics be i-maxims, a-maxims, or k-maxims? That is, does the Categorical Imperative (CI) assess principles of action as they are adopted by concrete individuals (A's particular lying maxim and B's particular lying maxim are assessed separately); does it assess principles of action independently of the individuals who adopt them (the maxim held in common by X and Y is assessed once, perhaps by appeal to the will of an ideal individual rather than by appeal to the respective wills of X and Y); or does it assess groups of maxims (e.g., lying promise maxims; the set of maxims adopted by X, Y, and Z)? In the remainder of this section, I shall explain this question in more detail.

The question has to do with logical priority. Here are three possibilities:

1. The locus of assessment is i-maxims: a given a-maxim or k-maxim is im/permissible because its corresponding i-maxims are (generally) so.
2. The locus of assessment is a-maxims: a given i-maxim is im/permissible because its corresponding a-maxim is so; a given k-maxim is im/permissible because its corresponding a-maxims are (generally) so.
3. The locus of assessment is k-maxims: a given i-maxim or a-maxim is im/permissible because it belongs to a k-maxim that is so.

These possibilities are obviously not exhaustive, but they suffice for my purposes.¹ In this paper, I intend to argue for 1. I shall argue for it in two ways. First, I shall argue that there is textual evidence in favor of 1. Then I shall argue that there are good philosophical grounds in favor of 1. Thus, I shall argue that there are reasons to think that Kant thought i-maxims are the locus of assessment and, further, that he was right about this. I shall conclude the paper by considering some objections.

¹Other possibilities are more pluralistic.

The problem space I am entering can be elucidated further by comparison with a debate about truth. Consider the following three sentences:

- A. Zwei Dinge erfüllen das Gemüth mit immer neuer und zunehmender Bewunderung und Ehrfurcht, je öfter und anhaltender sich das Nachdenken damit beschäftigt: der bestirnte Himmel über mir und das moralische Gesetz in mir.²
- B. Two things fill the mind with always newer and ever increasing admiration and awe, the oftener and more steadily reflection occupies itself therewith: the be-starred heaven above me and the moral law within me.³
- C. Two things fill the mind with always newer and ever increasing admiration and awe, the oftener and more steadily reflection occupies itself therewith: the be-starred heaven above me and the moral law within me.

One can use these three sentences to distinguish between sentence tokens (utterances), sentence types, and propositions. A, B, and C are distinct sentence tokens, but B and C are the same sentence type, and all three express the same proposition. Armed with this distinction, one might debate about whether sentence tokens, sentence types, or propositions are the primary truth-bearers. Hopefully the parallel is clear.

There is also a parallel with a well-known debate in utilitarianism: the debate between act and rule utilitarians. This debate is often explained as a debate about whether to apply the principle of utility to act tokens or to act types.⁴ To say that the principle of utility should be applied to act tokens (to advocate act utilitarianism) is analogous to saying that Kant's CI should be applied to i-maxims. To say that the principle of utility should be applied to act types (to advocate rule utilitarianism) is analogous to saying that Kant's CI should be applied to a-maxims. There is no analog, as far as I am aware, to saying that Kant's CI should be applied to k-maxims in debates about utilitarianism. But the purpose of this comparison is not to have a one-to-one mapping or a perfect analogy, nor is it to draw any conclusions about utilitarianism. Rather, the purpose is merely to try to make clearer what I am going to argue for in this paper.

There are debates among Kantians about the proper form of a well-formed maxim.⁵ There also are debates among Kantians about whether principles of omission or principles with the form 'to do X if I like' count as

²KpV, AA 05: 161.33–36.

³Although all translations in this paper are my own, I have consulted Gregor's.

⁴(Smart, 1956).

⁵For example, Rawls says that a maxim has this standard form: 'I am to do X in circumstances C in order to bring about Y unless Z' (Rawls, 2000, 168).

genuine maxims.⁶ And there are debates about how general a principle must be for it to count as a maxim (rather than a mere rule).⁷ But there has yet to be a debate about the difference between taking i-maxims, a-maxims, or k-maxims to be the locus of assessment.

Nonetheless, it is possible to extrapolate where some Kantians stand. For example, Rawls says that Kant's constructive procedure 'is thought of as embedding ... the ideal of a rational and reasonable person who understands and applies the categorical imperative procedure correctly.'⁸ From this it may be inferred (I shall return to this below) that Rawls (and many students of Rawls who follow his constructivist procedure) would disagree with me. Rawls does not take i-maxims to be the locus of assessment. Rather, Rawls takes a-maxims to be the locus of assessment.

I think that Rawls is wrong about this, and that is what I intend to show in this paper.

2. *The textual evidence*

I would like to introduce four pieces of textual evidence in favor of my thesis. The first is Kant's universalizability formulations of the CI.

The CI is supposed to encapsulate the Supreme Law of Morality. It is categorical because its bindingness is supposed to be unconditional: agents should follow the dictates of the moral law regardless of whether these dictates conflict with prudence or the pleasure of the moment.⁹

The CI is an imperative because Kant is interested in addressing human agents. Humans are subject to various limitations and hindrances. As a result of these limitations and hindrances, human behavior does not always conform to the Supreme Law of Morality. A holy will always acts in accordance with reason.¹⁰ But humans require constraint in the form of duty and respect, 'the representation of a worth, which derogates from my self-love.'¹¹ Thus, Kant presents the Supreme Law of Morality as an imperative.¹²

Kant articulates the CI in various forms.¹³ The first of the two so-called universalizability formulations is the Formula of Universal Law (FUL). It

⁶Wood denies that conditional principles of omission count as genuine maxims (Wood, 1999, p. 359 n28). Korsgaard denies that a principle like 'to do X if I like' counts as a genuine maxim (Korsgaard, 2013, p. 633).

⁷(Allison, 1990, chapter 5 section 1).

⁸(Rawls, 2000, p. 238 n2).

⁹GMS, AA 4: 414.12–25.

¹⁰GMS, AA 4: 414.1–11.

¹¹GMS, AA 4: 401.28–29.

¹²GMS, AA 4: 412.26–413.11.

¹³There is dispute about how many forms there are and about how they relate to each other. See, for example, Guyer (2000, chapter 5).

is introduced first in section I of the *Groundwork for a Metaphysics of Morals* and then again in section II. It runs as follows: 'act only according to that maxim, through which you at the same time can will, that it become a universal law.'¹⁴

The second of the two universalizability formulations is the Formula of a Law of Nature (FLN). It is introduced almost immediately after FUL in section II of the *Groundwork for a Metaphysics of Morals*. It runs as follows 'act thusly, as if the maxim of your action through your will should become a universal law of nature.'¹⁵

I would like to draw attention to the following fact: in both FUL and FLN, the maxim is supposed to be universalized through 'your' will.¹⁶ This renders the tests different from an imperative like 'conform your will to universal laws [of nature].'¹⁷ The difference is that I might be able to *will* some principle as a universal law [of nature] even though it cannot *be* such a law and vice versa (I might *not* be able to *will* some principle as a universal law [of nature] even though it can *be* such a law). Let me explain.

An agent's beliefs determine, to some extent, what she is able to will. This is hardest to deny and correspondingly easiest to see when the beliefs concern facts about the physical world that determine what is possible. For instance, a member of the flat Earth society is able to will various maxims that I am not. This can impinge on duty. Moreover, Kant's discussion of the highest good (HG) in the *Critique of the Power of Judgment* shows that he was aware of this.

The HG is a moral world in which agents are as happy as they are virtuous. Kant maintains that the duty to promote the HG generates a warrant for belief in God. He bridges the gap from the HG to belief in God with the assertion that the duty to promote the former can be satisfied only if God exists. But in the *Critique of the Power of Judgment*, Kant says that an agent who is unable to bring herself to believe in God (and, thus, the real possibility of the HG) should give up on this duty:

This proof, which one easily can adjust to the form of logical precision, does not intend to say: it is just as necessary to accept the existence of God as to recognize the validity of the moral law; hence, whoever cannot convince himself of the first, could judge himself to be free from the bindingness of the second. No! Only the intention of the final end, brought to pass through the

¹⁴GMS, AA 4: 421.7–8, emphases removed.

¹⁵GMS, AA 4: 421.18–20, emphases removed.

¹⁶In the German, Kant uses the informal du-form (*deiner* Handlung, *deine* Wille), which lends support to my position. I would like to thank an anonymous referee for *Pacific Philosophical Quarterly* for pointing this out to me.

¹⁷Wood and Aune argue that this represents a gap in Kant's derivations of the universalizability tests (Aune, 1980, p. 29 and Wood, 1999, pp. 78–79). This is also discussed in (Kahn, 2014, section 1), which makes a parallel point regarding Kant's so-called Formula of Humanity.

observance of the latter, in the world (a happiness of rational beings harmoniously corresponding with the observance of the moral law, as the highest and world's best) would have to be given up in that case.¹⁸

Kant's point in this passage is that, if someone is unable to convince herself that it is possible to fulfill a duty (in this case, the duty to promote the HG), this does not abrogate the bindingness of the moral law *tout court* but rather only the bindingness of that particular duty. His argument relies on a tacit appeal to the principle that ought implies can (OIC), something I shall return to below. But for now the point is that, according to Kant, (i) what an agent believes influences what she can will, and (ii) this (in conjunction with OIC) entails that different agents can have different duties.

The connection between beliefs and duties is more direct in the universalizability tests than it is in the HG argument. The connection between beliefs and duties arises in FUL and FLN because whether a maxim can become a universal law [of nature] through *my* will depends to some extent on *my* beliefs about contingent, empirical facts. For instance, whether I can will a lying promise maxim or a maxim of indifference as a universal law [of nature] depends to some extent on *my* beliefs about human memories and gullibility, conditions of relative scarcity, and so on.

I do not want to say that Kant was consciously committed to the idea that different agents can will different maxims as universal laws [of nature] and, thus, that Kant was consciously committed to the idea that the deontic status of a maxim can vary from agent to agent, although he does come close to saying as much in the HG argument passage above. I think that Kant *should* have been committed to this. But that is not the point. Rather, the point is to draw attention to the following fact: FUL and FLN are directed at individual agents' wills and, more, individual agents' maxims.

In other words, according to the universalizability tests, the permissibility of a given, concrete, individual maxim lies in the will of a given, concrete, individual agent. Even if the universalizability tests give qualitatively identical results for qualitatively identical i-maxims, the point is that the results can be generated only by ascertaining whether a given, concrete, individual agent can will the i-maxim as a universal law [of nature].¹⁹

Compare this with Rawls' constructivism. Rawls says that Kant should be thought of as presenting the principles 'from which a fully ideal reasonable and rational agent would act.'²⁰ Thus, on Rawls' account of the universalizability tests, the permissibility of a maxim is ascertained by determining whether an ideal agent, one 'endowed with conscience and moral

¹⁸KU, AA 5: 450.31–451.7.

¹⁹There is then a deep question, one that I cannot broach right now, about whether and how Kant himself avoids the conclusion that different agents generally can will different maxims as universal laws [of nature].

²⁰(Rawls, 2000, p. 151).

sensibility,' one endowed with various true beliefs about the world as well as psychological proclivities, could will the maxim as a universal law [of nature] 'in the normal conditions of human life.'²¹

Rawls takes the universalizability tests to assess maxims at a higher level of abstraction than i-maxims: a-maxims. The assessment is carried out by constructing an abstract, ideal agent who is considered to be in abstract, ideal ('normal') conditions. The question Rawls asks is whether this ideal agent can will a given a-maxim as a universal law [of nature] in these ideal conditions. Rawls then infers the deontic status of all corresponding i-maxims.

I want to argue, however, that this involves a subtle but important misconstrual of Kant. Rawls' account requires reading the universalizability tests as instructing non-ideal agents who might be in non-ideal conditions to act only on maxims that an ideal agent can will as universal laws [of nature] in ideal conditions. Or, if you prefer, Rawls' account requires reading the CI as telling agents to 'conform your will to universal laws [of nature].' The very care with which Rawls sets up his account evinces the fact that there is a difference between FUL and FLN as applied to a concrete agent's will (*your* will – what *you* can will as a universal law [of nature] given *your* beliefs and conditions) and FUL and FLN as applied to an ideal agent's will given ideal beliefs and ideal conditions, and I think that highlighting this difference militates in favor of my thesis.

I turn now to the second piece of textual evidence for my thesis: Kant's examples of how to apply FLN in section II of the *Groundwork for a Metaphysics of Morals*.

After articulating the universalizability tests, Kant illustrates their usefulness with four examples. The first involves an agent who 'feels a weariness in life' and contemplates suicide.²² The second involves an agent thinking about making a lying promise to repay when he 'sees himself pushed by necessity, to borrow money.'²³ The third involves an agent 'in comfortable circumstances' and thus deliberating about a life of pleasure notwithstanding the fact that he has 'a talent that by means of some cultivation could make him into a person useful for all sorts of intention.'²⁴ And the fourth involves an agent in comfortable circumstances who is weighing whether to let every individual 'be as happy as heaven wills, or [as] he can make himself' rather than to aid others in need.²⁵

What I would like to draw attention to is Kant's treatment of maxims in these examples. In particular, I would like to draw attention to the fact that

²¹(Rawls, 2000, p. 164).

²²GMS, AA 4: 421.24.

²³GMS, AA 4: 422.15.

²⁴GMS, AA 4: 422.37–423.2.

²⁵GMS, AA 4: 423.19–20.

the closest Kant gets to articulating the maxims to be tested in these examples is in the following text from the suicide case:

His maxim however is: I make it for me from self-love into a principle, when life with its longer duration threatens more badness than it promises amenity, to cut it short for myself.²⁶

Note that this text is muddled. At first blush, one would expect that everything after the colon is the agent's maxim. But in a footnote just two pages earlier, Kant remarks that a maxim is an agent's 'subjective principle of acting.'²⁷ This would lead one to think that the agent's maxim is what is in apposition with 'principle,' everything after the first comma. If this is correct, then self-love is the motive for adopting the maxim (not part of the maxim itself). Further inspection, however, leads back to the first-blush reading: Kant says that the reason this maxim fails FLN is that the motive of self-love contradicts the principle (of foreshortening life). So, which reading is correct? Is a maxim a principle, or is a maxim a principle and a motive? And what is a principle anyway?

I do not propose to answer these questions here. That is not why I am raising them (I return to this in Section 4). There are two reasons why I want to draw attention to these questions. The first is that I think Kant's lack of precision on this score speaks in favor of my thesis: Kant is more concerned with describing a plausible vignette of an actual, concrete agent deliberating about some i-maxim to which his readers can relate at some level than he is with establishing that some abstract a-maxim is im/missible in all of its individual instances.

The second is that I think the difference between Kant and modern discussions of Kant's examples on exactly this issue shows that the standard reading of Kant's universalizability tests, like Rawls', does not take them to apply to i-maxims. That is, in modern discussions of these four examples, commentators often take great care to articulate the following three things: (i) the form a maxim must have in general (so that a maxim is not well-formed if it lacks, for example, a description of an agent's purpose or the circumstances in which an action is to be performed); (ii) the a-maxim in question in each of Kant's examples; and (iii) the universalized typified counterpart of the maxim.²⁸ Depending on the interpretation of the universalizability tests, this is then used to give a rigorous demonstration of why the a-maxim in (ii) is im/missible, from which the deontic status of all individual instances of that maxim may be inferred.

I am arguing that the tests work in the other direction. The tests work by examining the i-maxim first and foremost; if I am right, then if conclusions

²⁶GMS, AA 4: 422.4–7.

²⁷GMS, AA 4: 420.36.

²⁸See, for example, Onora O'Neill's *Acting on Principle*, recently published in second edition by Cambridge University Press (O'Neill, 2013), or see Furer (2017a) and Furer (2017b).

are going to be drawn about a-maxims or k-maxims, these conclusions must be based on various generalizations about agents and their i-maxims. I shall return to this idea in the next section of this paper. For now, I turn to the third piece of evidence in favor of my thesis: Kant's commitment to OIC.²⁹

One reason for ascribing OIC to Kant was discussed earlier: in his HG argument in the *Critique of the Power of Judgment*, Kant appeals to OIC to make the following move: (1) the real possibility of the HG presupposes God and (2) there is a duty to promote the HG; therefore, (3) rational agents should bring themselves to believe in God or give up the duty to promote the HG.

A deeper reason for ascribing OIC to Kant follows from his conception of duty. According to Kant's conception of duty, an agent (*ultima facie*) ought to D only if (a) she would D were she governed by reason and (b) it is possible for her to govern herself in accordance with reason. From this, it may be seen that OIC is built into Kant's conception of duty.³⁰

OIC is relevant for present purposes because different individuals with different abilities will have different duties on the basis of this principle. Thus, qualitatively identical i-maxims might have different deontic statuses for different agents. So, Kant's commitment to OIC gives indirect, doctrinal support for the idea that i-maxims are the locus of assessment in Kant's ethics.

The fourth and final piece of evidence I would like to introduce in favor of my thesis is Kant's theory of conscience. In particular, I would like to draw attention to the following sentence from Kant's *Metaphysics of Morals*: 'when however someone is conscious of himself as having acted according to conscience, then of him, what concerns guilt or innocence, nothing more can be demanded.'³¹ This is notable because it comes only a few lines after Kant's claim that 'in the objective judgment, whether something is a duty or not, one can well err from time to time.'³² Putting these two sentences together it may be seen that, according to Kant, the deontic status of a maxim depends on whether it accords with an individual's conscience.

That is, the objective judgment of the deontic status of a maxim (whether a maxim actually could be a universal law [of nature]) might diverge from an agent's conscientious judgment about the deontic status of that maxim (whether the agent can will that maxim as a universal law [of nature]). When that happens, the 'ought' or 'normative bindingness' tracks the agent's conscientious judgment rather than the objective deontic status. This means

²⁹Kant's commitment to OIC is widely accepted. Indeed, OIC sometimes is referred to as 'the Kantian thesis,' 'the famous Kantian maxim,' or 'Kant's dictum' (Frankfurt, 1988, p. 95; Yaffe, 1999, p. 218; and Frankena, 1958, p. 60).

³⁰Kant also is committed to the claim that an agent has a duty to X only if she must be constrained to do so (see note 12 above and the paragraph to which it is appended). From this, it follows that Kant is committed to a principle related to OIC, ought-implies-able-not-to. I discuss this at greater length in my (Kahn, 2021b).

³¹MS, AA 6: 401.11–13.

³²MS, AA 6: 401.5–6.

(again) that qualitatively identical maxims might have different deontic statuses for different agents, and this (again) militates in favor of my thesis.

It might be objected that there is another way to accommodate this aspect of Kant's theory of conscience, one that does not require saying that the deontic status of qualitatively identical i-maxims can differ for different agents. In particular, it might be argued that an agent who makes a conscientious error has an excusing condition that renders her inculpable for doing something impermissible. This would enable one to hold onto the idea that the deontic status of the maxim in question is invariant across agents (regardless of their conscientious judgments).

But this accommodates only one side of Kant's theory of conscience. The other side of Kant's theory of conscience is that an agent who acts against her conscience does wrong. In the sentence immediately preceding, the one in which Kant says that acting in accordance with conscience is all that can be demanded of someone, Kant asserts: 'lack of conscience is not deficiency of conscience, but rather a propensity not to turn oneself to its judgments.'³³ On Kant's account, it is not only innocence but also guilt that tracks an agent's conscientious judgment.

I shall consider additional objections to my thesis below (in Section 4). For now, I want to argue that Kant was right to regard i-maxims as the locus of assessment. Accordingly, I turn to some philosophical arguments in favor of my thesis.

3. *The philosophical evidence*

It might be thought that there is no difference, philosophically speaking, between saying that the locus of assessment should be i-maxims as opposed to a-maxims or k-maxims. I think that this is incorrect. Not only do I think that there is a difference, philosophically speaking, between evaluating i-maxims as opposed to a-maxims or k-maxims, but, more, I think that there is a philosophical payoff to doing so. Accordingly, in this section I am going to canvass five arguments in favor of this.

First, I think that there is a metaphysical difference between evaluating i-maxims as opposed to other (more abstract) kinds of maxims. In particular, evaluating other a-maxims or k-maxims requires commitment to the existence of agent-independent principles (a-maxims) and agent-independent sets of the same (k-maxims). Given the anti-platonist leanings that seem to be popular among value theorists these days, I suggest that this is not only a difference but also an advantage.

Second, I think that the move toward i-maxims makes more explicit the fact that Kant's ethics need not be conceived as an absolutist set of duties

³³MS, AA 6: 401.10–11.

invariant across time, culture, and people. The CI is a core, objective standard that allows for variation by being applied to different agents' i-maxims, striking a favorable middle ground between the poles of ethical relativism and absolutism.³⁴ This, I grant, can be captured by the idea that an a-maxim is permissible-*for-an-agent*. But I suggest that, if the opponent of my thesis is arguing for applying the universalizability tests to an a-maxim to yield *this* result, then I am unsure what remains of the view that i-maxims are not the locus of assessment.³⁵

Third, if the locus of assessment is i-maxims, this gives bite to a traditional debate about highly specific maxims. That is, it is often argued that highly specific maxims, by virtue of being highly specific, always can be willed at the same time as universal laws [of nature] without contradiction, thus resulting in a host of false positives for the universalizability tests (i.e., intuitively impermissible maxims that the tests dub permissible). This is sometimes taken to be a fatal objection to the universalizability tests.³⁶

One popular response to this objection in the literature is that no actual agent ever adopts these maxims. But objectors are quick to point out that this is irrelevant: the universalizability tests nonetheless render these highly specific maxims permissible for all agents even if no agent actually adopts them, and that is taken to be counterintuitive.

But if i-maxims are the locus of assessment, then the fact that no actual agent ever adopts a highly specific maxim becomes a more tenable response. If the universalizability tests assess an individual agent's actual maxims to determine whether that individual can will them as universal laws [of nature], then it is not irrelevant to insist that no actual agent has adopted a given maxim or that no actual agent has adopted such a maxim without adopting a stack of other impermissible maxims.

Fourth, if the locus of assessment is i-maxims, this also gives new life to a traditional debate about the Kantian duty of beneficence. It sometimes is objected that the Rawlsian appeal to 'true needs' to support the duty of beneficence (Kant's fourth example described in Section 2) makes the results of the universalizability tests empirical in a way that they should not be. And, indeed, it is difficult to see why an ideal agent would have any 'true needs,' much less the same true needs as human agents.

But if i-maxims (rather than a-maxims) are being assessed, then this objection is ill-founded. The question to be asked is not whether some abstract, ideal agent with no (or few) needs can will a maxim of indifference as a

³⁴Along the same lines: the move to i-maxims can help put to rest an objection sometimes raised against Kantian ethics (inspired, I believe, from Bernard Williams), that the supposed monolithic nature of rationality entails that when someone is faced with a tough choice and nonetheless chooses rationally, agents who are on the losing side have no moral grounds for complaint.

³⁵One thing that would remain is the metaphysical baggage that comes along with opposing my thesis.

³⁶For a discussion, see section 7.3 of chapter 3 of Wood (1999).

universal law [of nature], nor is the question whether such a maxim is permissible in the abstract. Rather, the question is whether concrete agents with actual maxims, actual wills, and actual needs are able to will i-maxims of indifference as universal laws [of nature]. And the only way to determine this is to assess whether willing such maxims as universal laws [of nature] would render it impossible to fulfill some other ends that the agents in question are unable to foreswear.

Fifth and finally, appealing to i-maxims gives the most plausible account of the deontic status of k-maxims. To see why, consider an alternative view propounded by Korsgaard.

Korsgaard introduces a distinction between duties taken as general, provisionally universal, and universal. A duty is taken as universal if but only if it is taken to have no exceptions so that when an exception is found, the duty is discarded; a duty is taken as provisionally universal if but only if it is treated as defeasibly or presumptively universal until an exception is found, and when an exception is found the duty is reformulated accordingly; a duty is taken as general if but only if exceptions are expected, and '[i]f ... we encounter an exception, nothing happens.'³⁷ Korsgaard argues that duties are most plausibly regarded as provisionally universal.

I think that Korsgaard gets things wrong here. Insofar as she is putting this forth as a reading of Kant, it fails to take into consideration (i) the fact that in the *Metaphysics of Morals* Kant articulates casuistical questions regarding duties, questions that he seems to regard as genuinely open but nonetheless not requiring duty revision³⁸; (ii) the fact that the duties Kant articulates in the *Metaphysics of Morals* come as the conclusions of arguments, whence it follows that they could not be revised without revising the premises of these arguments³⁹; and (iii) the fact that Kant frequently says explicitly that a certain number of exceptions are expected for any general rule, and the only revision he hints at is that, if this indeterminate number is exceeded, then the rule will collapse.⁴⁰

Insofar as Korsgaard is putting this forth as an independently plausible philosophical thesis, I think it fails to take into consideration that this is not how morality works in practice: children are taught simple moral rules, like the rule not to lie, at a young age (think of the apocryphal story of George Washington and his cherry tree); they grow up and steadily come across more and more exception cases; but rather than teaching the next generation the more refined rules that account for these exception cases, the process is repeated.

But I want to argue that these mistakes spring from a deeper problem: a focus on maxims at the wrong level of abstraction. It makes no sense to talk

³⁷(Korsgaard, 1999, p. 24).

³⁸For example, MS, AA 6: 423.17–424.8.

³⁹For example, MS, AA 6: 421.10–423.16.

⁴⁰For example, LE, AA 27: 574.17–20.

of exceptions outside of the context of k-maxims, sets of maxims grouped by common content. Korsgaard is taking a top-down approach to duties.

If, by way of contrast, i-maxims are the locus of assessment, then the correct approach to inferring the deontic status of k-maxims is to begin with plausible generalizations about human nature and the environments in which humans tend to live, and then to use these generalizations to ground inferences to further generalizations about whether the i-maxims corresponding to a given a-maxim are going to pass the universalizability tests – that is, whether concrete individual agents in general will be able to will these maxims as universal laws [of nature]. It is hard to see how the universalizability tests could be applied to k-maxims directly, and I want to suggest that the kind of bottom-up approach to duties just outlined is more plausible intuitively and conforms more to the lived experience of morality than a top-down approach.

4. *Some objections*

I am going to wrap up my paper by considering five objections to my thesis. Although I think that none of the objections works, I think that they are useful inasmuch as they elucidate what is (and what is not) at stake. I use OI, OII, OIII, OIV, and OV to refer to the authors of the objections.

Objection I: the distinction between i-maxims, a-maxims, and k-maxims presupposes an answer to the question of what qualifies as a maxim. As OI puts the point, 'It is not even clear how the distinction between different types of maxims can be established without answering the question as to what qualifies as a maxim in the first place.'

There is a grain of truth to this objection. Without some minimal concept of what a maxim is, there is no way to determine whether it makes sense to talk about i-maxims, a-maxims, and k-maxims, let alone to ask which should be the locus of assessment in Kant's ethics.

However, we have such a minimal concept, and I articulated it in Section 1 of this paper. All sides concede that a maxim is a subjective principle of volition, and that is all that I need in order to motivate my thesis. What I do not need, and what this objection mistakenly thinks that I need, is to take a stand on the precise form of a maxim. My thesis is orthogonal to such debates. To see why this is so, consider the debate about whether maxims can be highly specific. As noted above, according to some, only very general principles count as genuine maxims; specific rules do not. There is further debate about how much generality is requisite. But the following two examples seem relatively uncontroversial: 'to subordinate self-love to the moral law' is sufficiently general on the view that only very general principles are genuine maxims, whereas 'to use the blue mug when I make tea' is not. The point that is relevant for present purposes is that philosophers on both sides of this

debate can make sense of and apply the distinction between i-maxims, a-maxims, and k-maxims. To see how, note that two agents who adopt the maxim ‘to subordinate self-love to the moral law’ have adopted distinct i-maxims but the same a-maxim. It follows immediately that the distinction between i-maxims, a-maxims, and k-maxims does not presuppose an answer to the question of what qualifies as a maxim, and thus, the objection does not withstand critical scrutiny.

Objection II: the distinction between i-maxims, a-maxims, and k-maxims presupposes an answer to the question of how the universalizability tests are supposed to work. In the words of OII, I need to say more about how I ‘understand the universalizability tests to work ... the discussion of contradictions in the will is left completely obscure, so we’re left to just guess at all that.’

As with objection I, there is a grain of truth to this objection. Without some idea of how the universalizability tests are supposed to work, these tests provide no guidance at all regarding the question I am asking in this paper.

However, and also as with objection I, we have some idea of how the universalizability tests are supposed to work. Both FUL and FLN are explicitly worded to be about whether a given agent is able to will her maxim in a certain way, and that is all I need in order to motivate my thesis. What I do not need, and what this objection mistakenly thinks that I need, is to take a stand on the precise understanding of how the universalizability tests are supposed to work. To see why I do not need this, consider the debate about whether the universalizability tests should be understood in terms of a practical contradiction or a logical contradiction. According to proponents of the practical contradiction interpretation, a maxim fails the so-called contradiction in conception test if an agent is unable to achieve her end using the means specified in her maxim in a world in which that maxim is a universal law of nature. According to proponents of the logical contradiction interpretation, by way of contrast, a maxim fails the contradiction in conception test if the means specified in her maxim no longer exist in a world in which that maxim is a universal law of nature. Thus, proponents of the practical contradiction interpretation point out that, if a lying-promise maxim is universalized, then promise-keeping no longer will be effective in attaining money (or whatever one’s desired ends might be), whereas proponents of the logical contradiction interpretation point out that, if a lying-promise maxim is universalized, then the promise-keeping convention no longer will exist. And the point now is that, as with objection I, philosophers on both sides of this debate can make sense of and apply the distinction between i-maxims, a-maxims, and k-maxims. To see how, note that it is open to a proponent of either interpretation to specify, as, indeed, Rawls does, an ideal agent who has only true beliefs and is in normal conditions, and to ask whether *that* agent would be able to conceive her maxim as a universal law of nature (whether a practical or logical contradiction would manifest) – or, alternatively, to ask, as I do, whether a given, concrete, individual agent is able to

conceive her maxim as a universal law of nature. In other words, we can ask whether the contradiction manifests in *the particular agent's* will (my thesis), or whether the contradiction manifests in the idealized will of an ideal agent (the contrary of my thesis) – and this question is independent of which interpretation is adopted regarding the kind of contradiction (practical, logical, teleological, etc.) meant to manifest in the universalizability tests. The distinction between i-maxims, a-maxims, and k-maxims does not presuppose an answer to the question of how the universalizability tests work, and nor does determining whether i-maxims, a-maxims, or k-maxims should be the locus of assessment for the universalizability tests. This objection does not withstand critical scrutiny.

Objection III: my thesis presupposes that the universalizability tests are free-standing norms. But this is mistaken. The universalizability tests can be applied only in conjunction with the other formulations of the CI and, indeed, perhaps only in conjunction with Kant's social and political philosophy as a whole as developed in the Doctrine of Right, especially the Universal Principle of Right. And these other parts of Kant's ethics (the other formulations of the CI and his Doctrine of Right) are inconsistent with my thesis. Thus, as OIII asserts, my thesis 'takes for granted that the Formula of Universal law is supposed to be a free-standing norm, the proper application of which is uninformed by any of the other formulas. That claim is hardly obvious (indeed, it seems false, because Kant clearly says that the FUL specifies only the 'Form' of moral judgments, but not yet its corresponding 'matter' (and so, not it's [sic] 'complete determination' with reference to a universal Kingdom of Ends).' According to OIII, the universalizability tests can be applied only in the context of responding to the actions of others, and 'responding to the actions of others brings us into Kant's broader social and political philosophy (and so, the notion of a moral right).'

I would like to say three things in response to this objection. First, my thesis does not presuppose that the universalizability tests are freestanding norms. Rather, one of the arguments I gave for my thesis (namely, the first textual argument in Section 2, which depends on the interpretation of the universalizability tests) presupposes that a maxim is permissible if but only if it passes the universalizability tests. But (a) presupposing that a maxim is permissible if but only if it passes the universalizability tests is not the same as presupposing that the universalizability tests are freestanding norms, and (b) it is not the case that, if an argument for X presupposes Y, then X presupposes Y. Moreover, (c) that a maxim is permissible if but only if it passes the universalizability tests is implicit in the way Kant employs the universalizability tests throughout his corpus, including his discussion of the four examples rehearsed in Section 2 of this paper; and (d) although I have not argued for it here (and do not have the space to do so at length), I believe that the other formulations of the CI, properly understood, either imply or, at the very least, are consistent with my thesis, and anybody who

thinks that, according to Kant, the different formulations of the CI are supposed to be equivalent, at least extensionally if not intensionally, should agree.

Second, it is not the case that the universalizability tests can be applied only in the context of responding to the actions of others. Such an assertion vitiates (and, in turn, is vitiated by) an entire category of duties that Kant makes central to his project: duties to oneself.⁴¹ But even if this assertion were true (it is not true, but even if it were), it would not imply that discussion of the CI and its locus of assessment necessitates discussion of Kant's Doctrine of Right, his social and political philosophy, or the Universal Principle of Right. All of these latter topics can and, I maintain, should be divorced from the former. Failure to understand this not only threatens to misunderstand all of Kant's practical philosophy, perhaps by amalgamating it with Hegel's idea of moving through the philosophy of right prior to ethics, but also would seem to require discarding as confused nearly every single article and book written on Kant's ethics in the last 240 years (most of which neglect the Doctrine of Right) and also Kant's own *Groundwork for a Metaphysics of Morals* and *Critique of Practical Reason*. I do not think this can form a sound foundation for an interpretation of Kant's ethics.

Third and finally, I see no reason why Kant's social and political philosophy, the Doctrine of Right, the Universal Principle of Right, or ethical duties that arise in the context of others would be even *prima facie* inconsistent with my thesis. This is not the place to try to articulate a full treatment of any of these topics; to do so would take us far afield. I note merely that I used examples of duties to others in Sections 2 and 3 to argue in favor of my thesis. Thus, it seems to me that the burden of proof lies firmly on the shoulders of my interlocutors to explain why this strategy is foredoomed. From all of this it may be inferred that I think this objection does not work and, more, that it should be rejected in the strongest terms on the grounds that it rests on fundamental misunderstandings of Kant's and Kantian ethics.

Objection IV: there is no functional difference between i-maxims and a-maxims from the first-personal perspective. Thus, from the first-personal perspective, my thesis is meaningless. Thus, OIV asserts, 'Kant's universalization tests are addressed to agents engaged in first-person practical deliberation. From such a perspective, there isn't any real difference between i-maxims and a-maxims, just as there's no functional difference to me in my normal reflections between 'I believe p' and 'p', despite their very different subject-matter and truth conditions.'

This objection rests on a misunderstanding of the difference between i-maxims and a-maxims. The difference between i-maxims and a-maxims

⁴¹Indeed, in the *Metaphysics of Morals* Kant goes so far as to say that 'if there were no duties to oneself, then there would be none at all, not even any external duties' (MM, AA 6: 427.24–25). I would like to thank an anonymous referee for *Pacific Philosophical Quarterly* for suggesting that I include this quotation to bolster the point I am making in sentence to which this note is appended.

has to do with levels of abstraction. An i-maxim is a particular, concrete maxim as adopted by a particular, concrete individual in particular, concrete circumstances in the same way that an act token is a particular, concrete action performed by a particular, concrete individual in particular, concrete circumstances and in the same way that an utterance is a particular, concrete speech-act performed by a particular, concrete individual in particular, concrete circumstances. An a-maxim is an abstract principle that multiple agents can adopt in the same way that an act type is an abstract action that multiple agents can perform and a sentence is an abstract speech-act that multiple agents can utter. I do not want to take a stand on Moore's paradox, and I do not need to: there is no analogy between 'I believe p' and 'p' on the one side and i-maxims and a-maxims on the other.

*Objection V: There is debate among Kant scholars about whether, according to Kant, agents always act according to maxims, or whether agents always ought to act according to maxims (but sometimes do not). One objection to the latter position is that, because non-maxim-based action falls outside the scope of the universalizability tests, it will make many actions ethically unassessable. As OV points out, Nyholm proposes that we can determine whether a non-maxim-based action is permissible 'by asking whether a person who was governed by a maxim that could at the same time be a universal law could perform the given act, on that basis, in the circumstances.'*⁴² And OV suggests, based on Nyholm's proposal, that, although I might think I do not need to take a stand on this debate, maybe I do, for Nyholm's proposal seems to be framed in terms of a-maxims.

I would like to say two things in response to this objection. First, I think that OV is correct that I cannot adopt Nyholm's proposal. But, that does not entail one or the other position in this debate. For example, a proponent of the position that agents sometimes do not act according to maxims might bite the bullet and say that, on Kant's account, non-maxim-based action is not directly ethically assessable. Depending on what kinds of actions end up not being maxim-based, this might not be such a big bullet to bite, especially because these actions still might be indirectly ethically assessable.

Second, I suspect that Nyholm's proposal might not work independently of the question I have addressed in this paper. The problem is that, on the one hand, there is a many-one relationship between a-maxims and actions and, on the other hand, Nyholm's proposal gives only a semi-decision procedure for determining whether an action is permissible. This means that we might never be able to figure out whether a given action is impermissible, and it also strongly suggests that there will be many false positives (impermissible actions that this proposal would label as permissible).⁴³

⁴²(Nyholm, 2017, 251).

⁴³I discuss some of these issues at greater length in my (Kahn, 2021a, section 2).

5. Conclusion

In this paper I have argued that Kant thought i-maxims are the locus of assessment and, further, that he was right about this. I examined four main pieces of evidence in favor of the textual side of my thesis: (i) the universalizability formulations of the CI; (ii) Kant's seeming unconcern with the exact formulation of the agents' maxims in his vignettes; (iii) Kant's commitment to OIC; and (iv) Kant's theory of conscience. I introduced five arguments in favor of the philosophical side of my thesis: (a) the metaphysical difference between my thesis and other positions in the problem space; (b) the way that my thesis allows for objectivity without falling into absolutism; (c) the problem of highly specific maxims; (d) the problem of true needs and beneficence; and (e) the deontic status of general duties. Finally, I considered five objections to my thesis: (I) my thesis makes no sense without a discussion of what qualifies as a maxim; (II) my thesis makes no sense without a discussion of how the universalizability tests work; (III) my thesis is built on an untenable view of the universalizability tests; (IV) there is no first-personal difference between i-maxims and a-maxims; and (V) I have to take a stand on the debate about whether agents always (ought to) act on maxims. I argued that none of these objections works.

Department of Philosophy
Indiana University-Purdue University, Indianapolis

REFERENCES

- Allison, H. (1990). *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Aune, B. (1980). *Kant's Theory of Morals*. Princeton: Princeton University Press.
- Frankena, W. (1958). 'Obligation and motivation in recent moral philosophy,' in A. I. Meldon (ed.) *Essays in Moral Philosophy*. Seattle: University of Washington Press, pp. 40–81.
- Frankfurt, H. (1988). *The Importance of What we Care about*. Cambridge: Cambridge University Press.
- Furner, J. (2017a). 'Kant's Contradiction in Conception Test,' *Theoria*, 64, No.3(152), pp. 1–23.
- Furner, J. (2017b). 'Kant's Contradiction in the Will Test,' *The Philosophical Forum* 48(3), pp. 307–323.
- Guyer, P. (2000). *Kant on Freedom, Law, and Happiness*. Cambridge: Cambridge University Press.
- Kahn, S. (2014). 'The Interconnection between Willing and Believing for Kant's and Kantian Ethics,' *International Philosophical Quarterly* 54(2), pp. 143–157.
- Kahn, S. (2021a). 'Obligatory Actions, Obligatory Maxims,' *Kantian Review* 26(1), pp. 1–25.
- Kahn, S. (2021b). 'Rethinking Kant on Duty,' *The Review of Metaphysics* 74, No. 4(296), pp. 497–526.
- Korsgaard, C. (1999). 'Self-Constitution in the Ethics of Plato and Kant,' *The Journal of Ethics* 3(1), pp. 1–29.
- Korsgaard, C. (2013). 'Kantian Ethics, Animals, and the Law,' *Oxford Journal of Legal Studies* 33(4.1), pp. 629–648.

- Nyholm, S. (2017). 'Do we Always Act on Maxims,' *Kantian Review* 22(2), pp. 233–255.
- O'Neill, O. (2013). *Acting on Principle*. Cambridge: Cambridge University Press.
- Rawls, J. (2000). *Lectures on the History of Moral Philosophy*. Cambridge: Harvard University Press.
- Smart, J. J. C. (1956). 'Extreme and Restricted Utilitarianism,' *The Philosophical Quarterly* 6(25), pp. 344–354.
- Wood, A. (1999). *Kant's Ethical Thought*. Cambridge: Cambridge University Press.
- Yaffe, G. (1999). "'Ought' Implies 'Can' and the Principle of Alternate Possibilities,' *Analysis* 59(1), pp. 218–222.